Algorithmic Differentiation for Climate Science **11/30/2015**

Sri Hari Krishna Narayanan
Argonne National Laboratory
Daniel Goldberg (University of Edinburgh), Patrick Heimbach (University of Texas at Austin), David Marshall (Oxford University), Stephen Price (LANL), Carl Wunsch (Harvard University)

Abstract

Computing sensitivities of climate-related quantities of interest (QoI) to very high-dimensional state or parameter spaces in an efficient manner is of considerable interest in climate science. Applications in which sensitivities are propagated in time by means of adjoint models include quantifying controlling factors of the QoI's variations [24, 3], optimal state and parameter estimation ("data assimilation") [25, 26], computing nonnormal transient amplification patterns [27], and formal uncertainty quantification [10, 11]. In an environment of rapid model development algorithmic differentiation (AD) plays a crucial role in supporting adjoint model generation for diverse applications based on up-to-date source codes [9]. With applications ranging from oceanography [19], atmospheric sciences [28], coupled carbon cycle modeling [12] and cryospheric sciences [5, 2], the need for sustained AD tool maintenance and improvements becomes a high priority.

Background/Research to Date

Climate models depend on a large (because spatially varying) parameter set that represents subgrid-scale processes and whose values are inherently uncertain. Similarly, initial and boundary conditions (at the surface or base) are uncertain and will dictate the solution. To understand model sensitivities derivatives are needed for relating the outputs of a model back to its inputs, i.e. control variables. Model calibration, in particular, which seeks to determine optimal estimates of control variables informed by available observations, plays an increasingly important role in attempts to initialize models for prediction. The conventional approach of finite-difference approximation becomes intractable for high-dimensional control spaces and simulations which require days for single-model execution on today's supercomputers. Furthermore, this approach can be inaccurate because of accumulating round-off errors.

An alternative method is the use of adjoint models. For scalar-valued QoI's such as least-squares misfit functions used in model calibration or state estimation, or for sensitivity analysis of climate metrics the adjoint provides an extremely efficient means for computing linear sensitivities. In climate modeling, the challenge, then, is to maintain up-to-date versions of the adjoint model given the evolving nature of the climate model at hand. Writing the adjoint by hand has frequently shown to be cumbersome, and producing code that is increasingly out of date with respect to its parent code. It also is prone to additional coding error.

AD provides a powerful alternative [7]. It exploits the code's composition of elementary arithmetic operations and elementary function evaluations whose analytical derivatives are known. Applying the chain rule of differentiation, derivatives can be propagated from one model variable to another. Various AD tools have been developed based on techniques from source transformation [21, 8, 4, 16] or operator-overloading [22, 17]. Current AD research is directed in part at increasing the efficiency of derivative code for components of large-scale climate models. Reverse-mode AD necessarily uses checkpointing where the application state is determined by source analysis. Many checkpointing schemes are being examined to efficiently use the available memory and disk while computing adjoints. Preliminary research allows the suspension and restart of the forward and the adjoint model. Adjoinable MPI seeks to differentiate through the MPI calls in the model code. AD advances are exploiting the inherent mathematical properties of the model to allow larger adjoint models to be run. For example, through the special treatment of fixed-point iterations, which differs from a straightforward application of AD software, both computational and memory requirements of the adjoint model are reduced.

Proposed Direction of work

We propose to extend the use and improve the efficiency of adjoint codes of climate model components. At the center of the effort is the open-source AD tool OpenAD [21].

• **New Languages** Many existing models are written in Fortran. As models migrate to C, C++ [14, 20], R, and Python, we propose to apply existing source transformation tools such as ADIC [16] for models written in C. Source transformation AD is not possible for C++ because of features such as templates and overloading. Instead, we propose to apply a mixed method [13] that uses operator overloading AD for most of the application and source transformation AD for the computation-intensive portions that are often written in a C-like fashion.

• **Scaling** An important bottleneck for scaling of adjoint computations is the memory requirement for storing partials and checkpointing of data at the granularity of timesteps. We will apply optimal multilevel checkpointing involving disk storage and memory [1, 18].

• **Robustness** The mean time between failure (MTBF) of leadership-class machines is on the order of hours. Because adjoint computations will often exceed MTBF, we propose to support the restart of the adjoint climate application using existing optimal checkpointing techniques in AD [6]. We will study the implementation of the restart mechanism in the context of multilevel checkpointing.

• **Chaotic Systems** Because of the increasingly nonlinear dynamics with time encountered in climate simulations, the adjoint results can diverge exponentially from the "macroscopic climate sensitivity" [15]. By replacing the initial value problem in these models with the well-conditioned "least squares shadowing (LSS) problem", it may be possible to compute well-behaved derivatives of climate quantities. [23]. We propose to study the non-straightforward application of AD tools for the LSS approach.

• **Usability and Applicability** Many climate modelers have been reluctant to employ AD. We will therefore strive to make source transformation tools easily applicable to climate codes. Moreover, we will work to demonstrate to the climate community the benefits of AD.

Connections to Math, Comp Sci & and Climate Science

The proposed work covers various fields, including, the solution of (discretized) PDEs, high-performance computing (techniques such as checkpointing and resilience) and different components of climate models (e.g. ice sheets, oceans, atmosphere, carbon cycle).

Potential Impact on the field

Climate modelers who already use AD will be able to run adjoint models at larger scales than they can currently run. Adjoint models that previously could not be executed because of machine failures can be executed with checkpointing support. Climate modelers who are hesitant to migrate to new languages will be able to do so more easily. Success stories may encourage other climate researchers to begin to use AD.

References

[1] G. Aupy, J. Herrmann, P. Hovland, and Y. Robert. *Optimal Multistage Algorithm for Adjoint Computation*. PhD thesis, LIP-ENS Lyon; Argonne National Laboratory, 2014.

[2] I. Fenty, D. Menemenlis, and H. Zhang. Global coupled sea ice-ocean state estimation. *Clim. Dyn.*, pages 1–26, Aug. 2015.

[3] I. Fukumori, O. Wang, W. Llovel, I. Fenty, and G. Forget. A near-uniform fluctuation of ocean bottom pressure and sea level across the deep ocean basins of the Arctic Ocean and the Nordic Seas. *Prog. Oceanogr.*, 134(C):152–172, May 2015.

[4] R. Giering and T. Kaminski. Applying TAF to generate efficient derivative code of Fortran 77-95 programs. In *Proceedings of GAMM 2002, Augsburg, Germany*, 2002.

[5] D. Goldberg, P. Heimbach, I. Joughin, and B. Smith. Committed near-future retreat of Smith, Pope, and Kohler Glaciers inferred by transient model calibration. *The Cryosphere Discuss.*, 9(4):4459–4498, 2015.

[6] A. Griewank and A. Walther. Revolve: An implementation of checkpointing for the reverse or adjoint mode of computational differentiation. *ACM Trans. Math. Softw.*, 26(1):19–45, 2000.

[7] A. Griewank and A. Walther. *Evaluating derivatives. Principles and techniques of algorithmic differentiation. 2nd ed.* SIAM., 2008.

[8] L. Hascoet and V. Pascual. The Tapenade automatic differentiation tool: Principles, model, and specification. *ACM Trans. Math. Softw.*, 39(3):20:1–20:43, May 2013.

[9] P. Heimbach, C. Hill, and R. Giering. An efficient exact adjoint of the parallel MIT General Circulation Model, generated via automatic differentiation. *Future Generation Computer Systems*, 21(8):1356–1371, Oct. 2005.

[10] A. G. Kalmikov and P. Heimbach. A Hessian-Based Method for Uncertainty Quantification in Global Ocean State Estimation. *SIAM J. Sci. Comput.*, 36(5):S267–S295, Sept. 2014.

[11] T. Kaminski, F. Kauker, H. Eicken, and M. Karcher. Exploring the utility of quantitative network design in evaluating Arctic sea ice thickness sampling strategies. *The Cryosphere*, 9(4):1721–1733, 2015.

[12] T. Kaminski, W. Knorr, G. Schürmann, M. Scholze, P. J. Rayner, S. Zaehle, S. Blessing, W. Dorigo, V. Gayler, R. Giering, N. Gobron, J. P. Grant, M. Heimann, A. Hooker-Stroud, S. HOUWELING, T. Kato, J. Kattge, D. Kelley, S. Kemp, E. N. Koffi, C. Köstler, P. P. Mathieu, B. Pinty, C. H. Reick, C. Rödenbeck, R. Schnur, K. Scipal, C. Sebald, T. Stacke, A. T. van Scheltinga, M. Voßbeck, H. Widmann, and T. Ziehn. The BETHY/JSBACH Carbon Cycle Data Assimilation System: experiences and challenges. *J. Geophys. Res. Biogeosci.*, 118(4):1414–1426, Oct. 2013.

[13] K. Kulshreshtha, S. Narayanan, and T. Albring. A mixed approach to adjoint computation with algorithmic differentiation. *IFIP Advances in Information and Communication Technology (IFIP AICT Series): System Modeling and Optimization*, Under Submission.

[14] E. Larour, H. Seroussi, M. Morlighem, and E. Rignot. Continental scale, high order, high spatial resolution, ice sheet modeling using the Ice Sheet System Model (ISSM). *Journal of Geophysical Research: Earth Surface*, 117(F1):n/a–n/a, 2012. F01022.

[15] D. J. Lea, M. R. Allen, and T. W. Haine. Sensitivity analysis of the climate of a chaotic system. *Tellus A*, 52(5), 2000.

[16] S. H. K. Narayanan, B. Norris, and B. Winnicka. ADIC2: Development of a component source transformation system for differentiating C and C++ . *Procedia Computer Science*, 1(1):1845–1853, 2010. ICCS 2010.

[17] E. T. Phipps, R. A. Bartlett, D. M. Gay, and R. J. Hoekstra. Large-scale transient sensitivity analysis of a radiation-damaged bipolar junction transistor via automatic differentiation. In C. H. Bischof, H. M. Bücker, P. D. Hovland, U. Naumann, and J. Utke, editors, *Advances in Automatic Differentiation*, pages 351–362. Springer, 2008.

[18] M. Schanen, O. Marin, H. Zhang, and M. Anitescu. Asynchronous two-level checkpointing scheme for large-scale adjoints in the spectral-element solver Nek5000. Technical Report ANL/MCS-P5422-1015, Mathematics and Computer Science Division, Argonne National Laboratory, 2015.

[19] D. Stammer, M. Balmaseda, P. Heimbach, A. Köhl, and A. Weaver. Ocean Data Assimilation in Support of Climate Applications: Status and Perspectives. *Annu. Rev. Marine Sci.*, 8(1):annurev–marine–122414–034113–28, 2016.

[20] I. K. Tezaur, M. Perego, A. G. Salinger, R. S. Tuminaro, and S. F. Price. Albany/FELIX: a parallel, scalable and robust, finite element, first-order Stokes approximation ice sheet solver built for advanced analysis. *Geoscientific Model Development*, 8(4):1197–1220, 2015.

[21] J. Utke, U. Naumann, M. Fagan, et al. OpenAD/F: A modular open-source tool for automatic differentiation of Fortran codes. *ACM Trans. Math. Softw.*, 34(4):18:1–18:36, 2008.

[22] A. Walther and A. Griewank. Getting started with ADOL-C. In U. Naumann and O. Schenk, editors, *Combinatorial Scientific Computing*. Chapman-Hall, 2012.

[23] Q. Wang, R. Hu, and P. Blonigan. Least squares shadowing sensitivity analysis of chaotic limit cycle oscillations. *Journal of Computational Physics*, 267:210 – 224, 2014.

[24] C. Wilson, K. J. Horsburgh, J. Williams, J. Flowerdew, and L. Zanna. Tide-surge adjoint modeling: A new technique to understand forecast uncertainty. *J. Geophys. Res. Oceans*, 118(10):5092–5108, Oct. 2013.

[25] C. Wunsch and P. Heimbach. Practical global oceanic state estimation. *Physica D*, 230(1-2):197–208, June 2007.

[26] C. Wunsch and P. Heimbach. Dynamically and Kinematically Consistent Global Ocean Circulation and Ice State Estimates. In *Ocean Circulation and Climate: A 21st Century Perspective*, pages 553–579. Elsevier Ltd., 2013.

[27] L. Zanna, P. Heimbach, A. M. Moore, and E. Tziperman. Upper-ocean singular vectors of the North Atlantic climate with implications for linear predictability and variability. *Q. J. R. Meteorol. Soc.*, 138(663):500–513, Jan. 2012.

[28] X. Zhang, X.-Y. Huang, and N. Pan. Development of the Upgraded Tangent Linear and Adjoint of the Weather Research and Forecasting (WRF) Model. *J. Atmos. Oceanic Technol.*, 30(6):1180–1188, June 2013.

Acknolwedgement

This work was supported in part by a grant from the U.S. Department of Energy, Office of Science, under contract DE-AC02-06CH11357.